DOCUMENT RESUME

ED 390 030 CS 012 319

AUTHOR Stahl, Steven A.; And Others
TITLE What Happens When Students Read Multiple Source Documents in History? Reading Research Report No. 45.
INSTITUTION National Reading Research Center, Athens, GA.; National Reading Research Center, College Park, MD.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 95
CONTRACT 117A20007
NOTE 49p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Controversial Issues (Course Content); *Critical Reading; High Schools; *High School Students; *Information Sources; Notetaking; *Reader Text Relationship; Reading Processes; Reading Research; United States History
IDENTIFIERS *Text Factors

ABSTRACT

Some educators (Ravitch, 1990) have suggested that students use multiple source documents to study history. Such documents could be primary sources, such as Congressional bills or eyewitness accounts, or secondary sources, such as later commentaries. This study examined the processes used when 19 tenth-grade high school students were presented source documents about a controversial incident in U.S. history, the Tonkin Gulf Incident and its aftermath, and asked to read these texts, either to describe the incident or the Senate action on the Tonkin Gulf Resolution, or develop an opinion about the incident or resolution. Results indicated that students did gain in the consistency of their mental models after reading at least two documents, but did not make any further gains after that. When compared to lay experts, they failed to make any growth after a first reading. Examination of their notes also indicated that students tended to take literal notes, regardless of the final task. Findings suggest that the students were using the initial readings to garner the facts about the incident or the resolution. If students were asked for a description, they tended to stay close to the text. If asked for an opinion, however, they tended to ignore the information in the texts they read, even though they may have taken copious notes. Findings also suggest that high school students may not be able to profit from multiple texts, especially those presenting conflicting opinions, without some additional instruction. (Contains 32 references, 2 notes, 5 tables, and 2 figures of data. An appendix of data is attached.) (Author/RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 390 030

What Happens When Students Read Multiple Source Documents in History?

Steven A. Stahl
Cynthia R. Hynd
Bruce K. Britton
Mary M. McNish
University of Georgia

Dennis Bosquet
Clarke County School District

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS 012 317

NRRC
National
Reading Research
Center

READING RESEARCH REPORT NO. 45
Fall 1995

BEST COPY AVAILABLE

NRRC

National Reading Research Center

What Happens When Students Read Multiple Source Documents in History?

Steven A. Stahl
Cynthia R. Hynd
Bruce K. Britton
Mary M. McNish
University of Georgia

Dennis Bosquet
Clarke County School District

READING RESEARCH REPORT NO. 45
Fall 1995

The work reported herein is a National Reading Research Project of the University of Georgia and University of Maryland. It was supported under the Educational Research and Development Centers Program (PR/AWARD NO. 117A20007) as administered by the Office of Educational Research and Improvement, U.S. Department of Education. The findings and opinions expressed here do not necessarily reflect the position or policies of the National Reading Research Center, the Office of Educational Research and Improvement, or the U.S. Department of Education.

NRRC National Reading Research Center

Executive Committee
Donna E. Alvermann, Co-Director
University of Georgia
John T. Guthrie, Co-Director
University of Maryland College Park
James F. Baumann, Associate Director
University of Georgia
Patricia S. Koskinen, Associate Director
University of Maryland College Park
Nancy B. Mizelle, Acting Associate Director
University of Georgia
Jamie Lynn Metsala, Interim Associate Director
University of Maryland College Park
Penny Oldfather
University of Georgia
John F. O'Flahavan
University of Maryland College Park
James V. Hoffman
University of Texas at Austin
Cynthia R. Hynd
University of Georgia
Robert Serpell
University of Maryland Baltimore County
Betty Shockley
Clarke County School District, Athens, Georgia
Linda DeGroff
University of Georgia

Publications Editors

Research Reports and Perspectives
Linda DeGroff, Editor
University of Georgia
James V. Hoffman, Associate Editor
University of Texas at Austin
Mariam Jean Dreher, Associate Editor
University of Maryland College Park
Instructional Resources
Lee Galda, *University of Georgia*
Research Highlights
William G. Holliday
University of Maryland College Park
Policy Briefs
James V. Hoffman
University of Texas at Austin
Videos
Shawn M. Glynn, *University of Georgia*

NRRC Staff
Barbara F. Howard, Office Manager
Kathy B. Davis, Senior Secretary
University of Georgia

Barbara A. Neitzey, Administrative Assistant
Valerie Tyra, Accountant
University of Maryland College Park

National Advisory Board
Phyllis W. Aldrich
Saratoga Warren Board of Cooperative Educational Services, Saratoga Springs, New York
Arthur N. Applebee
State University of New York, Albany
Ronald S. Brandt
Association for Supervision and Curriculum Development
Marshá T. DeLain
Delaware Department of Public Instruction
Carl A. Grant
University of Wisconsin-Madison
Walter Kintsch
University of Colorado at Boulder
Robert L. Linn
University of Colorado at Boulder
Luis C. Moll
University of Arizona
Carol M. Santa
School District No. 5
Kalispell, Montana
Anne P. Sweet
Office of Educational Research and Improvement, U.S. Department of Education
Louise Cherry Wilkinson
Rutgers University

Production Editor
Katherine P. Hutchison
University of Georgia

Dissemination Coordinator
Jordana E. Rich
University of Georgia

Text Formatter
Ann Marie Vanstone
University of Georgia

NRRC - University of Georgia
318 Aderhold
University of Georgia
Athens, Georgia 30602-7125
(706) 542-3674 Fax: (706) 542-3678
INTERNET: NRRC@uga.cc.uga.edu

NRRC - University of Maryland College Park
3216 J. M. Patterson Building
University of Maryland
College Park, Maryland 20742
(301) 405-8035 Fax: (301) 314-9625
INTERNET: NRRC@umail.umd.edu

About the National Reading Research Center

The National Reading Research Center (NRRC) is funded by the Office of Educational Research and Improvement of the U.S. Department of Education to conduct research on reading and reading instruction. The NRRC is operated by a consortium of the University of Georgia and the University of Maryland College Park in collaboration with researchers at several institutions nationwide.

The NRRC's mission is to discover and document those conditions in homes, schools, and communities that encourage children to become skilled, enthusiastic, lifelong readers. NRRC researchers are committed to advancing the development of instructional programs sensitive to the cognitive, sociocultural, and motivational factors that affect children's success in reading. NRRC researchers from a variety of disciplines conduct studies with teachers and students from widely diverse cultural and socioeconomic backgrounds in pre-kindergarten through grade 12 classrooms. Research projects deal with the influence of family and family-school interactions on the development of literacy; the interaction of sociocultural factors and motivation to read; the impact of literature-based reading programs on reading achievement; the effects of reading strategies instruction on comprehension and critical thinking in literature, science, and history; the influence of innovative group participation structures on motivation and learning; the potential of computer technology to enhance literacy; and the development of methods and standards for alternative literacy assessments.

The NRRC is further committed to the participation of teachers as full partners in its research. A better understanding of how teachers view the development of literacy, how they use knowledge from research, and how they approach change in the classroom is crucial to improving instruction. To further this understanding, the NRRC conducts school-based research in which teachers explore their own philosophical and pedagogical orientations and trace their professional growth.

Dissemination is an important feature of NRRC activities. Information on NRRC research appears in several formats. *Research Reports* communicate the results of original research or synthesize the findings of several lines of inquiry. They are written primarily for researchers studying various areas of reading and reading instruction. The *Perspective Series* presents a wide range of publications, from calls for research and commentary on research and practice to first-person accounts of experiences in schools. *Instructional Resources* include curriculum materials, instructional guides, and materials for professional growth, designed primarily for teachers.

For more information about the NRRC's research projects and other activities, or to have your name added to the mailing list, please contact:

Donna E. Alvermann, Co-Director
National Reading Research Center
318 Aderhold Hall
University of Georgia
Athens, GA 30602-7125
(706) 542-3674

John T. Guthrie, Co-Director
National Reading Research Center
3216 J. M. Patterson Building
University of Maryland
College Park, MD 20742
(301) 405-8035

NRRC Editorial Review Board

Peter Afflerbach
University of Maryland College Park

Jane Agee
University of Georgia

JoBeth Allen
University of Georgia

Janice F. Almasi
University of Buffalo–SUNY

Patty Anders
University of Arizona

Harriette Arrington
University of Kentucky

Marlia Banning
University of Utah

Jill Bartoli
Elizabethtown College

Janet Benton
Bowling Green, Kentucky

Irene Blum
Pine Springs Elementary School
Falls Church, Virginia

David Bloome
Amherst College

John Borkowski
Notre Dame University

Karen Bromley
Binghamton University

Martha Carr
University of Georgia

Suzanne Clewell
Montgomery County Public Schools
Rockville, Maryland

Joan Coley
Western Maryland College

Michelle Commeyras
University of Georgia

Linda Cooper
Shaker Heights City Schools
Shaker Heights, Ohio

Karen Costello
Connecticut Department of Education
Hartford, Connecticut

Jim Cunningham
Gibsonville, North Carolina

Karin Dahl
Ohio State University

Marcia Delany
Wilkes County Public Schools
Washington, Georgia

Lynne Diaz-Rico
California State University-San Bernardino

Ann Egan-Robertson
Amherst College

Jim Flood
San Diego State University

Dana Fox
University of Arizona

Linda Gambrell
University of Maryland College Park

Mary Graham
McLean, Virginia

Rachel Grant
University of Maryland College Park

Barbara Guzzetti
Arizona State University

Frances Hancock
Concordia College of Saint Paul, Minnesota

Kathleen Heubach
University of Georgia

Sally Hudson-Ross
University of Georgia

Cynthia Hynd
University of Georgia

Gay Ivey
University of Georgia

David Jardine
University of Calgary

Robert Jimenez
University of Oregon

Michelle Kelly
University of Utah

James King
University of South Florida

Kate Kirby
Gwinnett County Public Schools
Lawrenceville, Georgia

Linda Labbo
University of Georgia

Michael Law
University of Georgia

Donald T. Leu
Syracuse University

Susan Lytle
University of Pennsylvania

Bert Mangino
Las Vegas, Nevada

Susan Mazzoni
Baltimore, Maryland

Ann Dacey McCann
University of Maryland College Park

Sarah McCarthey
University of Texas at Austin

Veda McClain
University of Georgia

Lisa McFalls
University of Georgia

Randy McGinnis
University of Maryland

Mike McKenna
Georgia Southern University

Barbara Michalove
Fowler Drive Elementary School
Athens, Georgia

Elizabeth B. Moje
University of Utah

Lesley Morrow
Rutgers University

Bruce Murray
University of Georgia

Susan Neuman
Temple University

John O'Flahavan
University of Maryland College Park

Marilyn Ohlhausen-McKinney
University of Nevada

Penny Oldfather
University of Georgia

Barbara M. Palmer
Mount Saint Mary's College

Stephen Phelps
Buffalo State College

Mike Pickle
Georgia Southern University

Amber T. Prince
Berry College

Gaoyin Qian
Lehman College-CUNY

Tom Reeves
University of Georgia

Lenore Ringler
New York University

Mary Roe
University of Delaware

Nadeen T. Ruiz
California State University-
Sacramento

Olivia Saracho
University of Maryland College Park

Paula Schwanenflugel
University of Georgia

Robert Serpell
University of Maryland Baltimore
County

Betty Shockley
Fowler Drive Elementary School
Athens, Georgia

Wayne H. Slater
University of Maryland College Park

Margaret Smith
Las Vegas, Nevada

Susan Sonnenschein
University of Maryland Baltimore
County

Bernard Spodek
University of Illinois

Steve Stahl
University of Georgia

Roger Stewart
University of Wyoming

Anne P. Sweet
Office of Educational Research
and Improvement

Louise Tomlinson
University of Georgia

Bruce VanSledright
University of Maryland College Park

Barbara Walker
Eastern Montana University-Billings

Louise Waynant
Prince George's County Schools
Upper Marlboro, Maryland

Dera Weaver
Athens Academy
Athens, Georgia

Jane West
Agnes Scott College

Renee Weisburg
Elkins Park, Pennsylvania

Allen Wigfield
University of Maryland College Park

Shelley Wong
University of Maryland College Park

Josephine Peyton Young
University of Georgia

Hallic Yupp
California State University

About the Authors

Steven A. Stahl is Professor of Reading Education at the University of Georgia and a principal investigator at the National Reading Research Center. He was formerly at the Center for the Study of Reading at the University of Illinois. After teaching reading in elementary schools in New York and Maine, he received his doctorate in reading from the Harvard Graduate School of Education. His research interests are centered around issues in reading instruction at the primary grades and the relationship between vocabulary knowledge and reading comprehension.

Cynthia R. Hynd is Associate Professor of Developmental Studies at the University of Georgia and a principal investigator with the National Reading Research Center. She received her doctorate in reading education from the University of Georgia. Her eight years of public school teaching included service as a remedial reading specialist and special education teacher. Dr. Hynd's research focuses on the cognitive aspects of learning from text. Her special interest is how students read textbooks in the sciences and social sciences.

Bruce K. Britton is a Professor of Psychology at the University of Georgia. He is well known for his work on understanding expository text. He has published six books, most recently *Learning from Textbooks: Theory and Practice*, and numerous journal articles in journals such as the *Journal of Educational Psychology*, *Discourse Processes*, and the *Journal of Experimental Psychology: Human Learning and Memory*.

Mary M. McNish is a doctoral student in the School Psychology Program at the University of Georgia. Currently, she is working on an internship at Grady Hospital in Atlanta

Dennis Bosquet is a history teacher at Clarke Central High School in Athens, GA.

National Reading Research Center
Universities of Georgia and Maryland
Reading Research Report No. 45
Fall 1995

What Happens When Students Read Multiple Source Documents in History?

Steven A. Stahl
Cynthia R. Hynd
Bruce K. Britton
Mary M. McNish
University of Georgia

Dennis Bosquet
Clarke County School District

Abstract. *Some educators (Ravitch, 1990) have suggested that students use multiple source documents to study history. Such documents could be primary sources, such as Congressional bills or eyewitness accounts, or secondary sources, such as later commentaries. This study examined the processes used when 19 high school students were presented source documents about a controversial incident in U.S. History, the Tonkin Gulf Incident and its aftermath, and asked to read these texts, either to describe the incident or the Senate action on the Tonkin Gulf Resolution, or develop an opinion about the incident or resolution. We found that students did gain in the consistency of their mental models after reading at least two documents, but did not make any further gains after that. When compared to lay experts, they failed to make any growth after a first reading. Examining their notes, we found that students tended to take literal notes, regardless of the final task. This suggests that they were using the initial readings to garner the facts about the incident or the resolution. If students were asked for a description, they tended to stay close to the text. If asked for an opinion, however, they tended to ignore the information in the texts they read, even though they may have taken copious notes. Our observations suggest that high school students may not be able to profit from multiple texts, especially those presenting conflicting opinions, without some additional instruction.*

After many years of comparative neglect, the study of history has received renewed attention by cognitive psychologists (Wineburg, 1991a, 1991b). Cognitive analyses of history learning have appeared in symposiums presented at major national meetings, as well as in books devoted to the subject (Leinhardt, Beck, & Stainton, 1994; Perfetti, Britt, & Georgi, 1995) and a special issue of *Educational Psychologist* (Wineburg, 1994).

This renewed attention may presage an interest in new methods of presenting historical content. The traditional means of teaching history was to rely heavily, if not exclusively, on the textbook as a means of conveying

information. In 1982, a survey found that roughly 90% of all social studies classes use a textbook in their class (Patrick & Hawke, 1982). Approximately half of all teachers in this survey reported relying on only one text, with that text reported as the major determinant of the content of their curriculum.

Currently, the single text approach to history learning and the model of learning upon which it is based are being challenged by both constructivist views of knowledge acquisition (Seixas, 1993) and more traditional views of history (Ravitch, 1992). This report examines an alternative approach to learning about historical events, using multiple original source materials, and the processes used by students as they negotiate this new approach.

Construction of Meaning in History

The view of the textbook-based teacher can be caricatured as a "transmission" model of learning: the information to be learned is contained in one vessel, the textbook, and is "transmitted" to another vessel, the student's memory, via the teacher's lecture. Traditionally, many teachers have treated content area knowledge as Hirsch (1987) did, as a "basket of facts" that must be gathered from text and lecture. These facts are stored in memory, just as one adds information to a computer database. As one history teacher put it, "History is the basic facts of what happened. What *did* happen. You don't ask how it happened. You just ask, 'What are the events?'" (Wineburg, 1991b, p. 513).

Such a transmission model is not supported by current views of the nature of knowledge and learning. More recent theories suggest that as information is learned, it is not copied merely from one source to another, but is transformed by the process of learning (Spiro, 1980). In this "constructivist" view of knowledge acquisition, new information can be retained in short-term memory through rote memorization or rehearsal, but this information is easily forgotten. This is evidenced by the often-experienced phenomenon of a student learning facts for a test and forgetting them as soon as the test is over. For information to be learned and retained, it must be combined actively with previously learned information. The new learning is "constructed" from the new information and the old information into new knowledge, either by *assimilating* the new knowledge into already existing knowledge structures or *accommodating* the new information by creating new knowledge structures that account for both the previously known and the new information (Rumelhart, 1980). Because every learner brings somewhat different knowledge and experience to the classroom, the knowledge that each learner retains is going to be somewhat different.

In this constructivist view of knowledge, the conveyance of content is more than merely insuring that the student devotes enough time and attention to memorizing the text or the teacher's lecture. Instead, the teacher must create the conditions that best allow the student to construct a mental model of the knowledge domain, incorporating not only the information in the current curriculum, but also past knowledge.

The constructivist view of learning not only challenges the transmission model of

learning, but also calls into question the relevance of those psychological models of learning based on the reading of a single text for examining the processes involved in learning history. Models such as Kintsch and van Dijk's (1978) model may accurately describe how readers construct a propositional text base from the reading of a single text. However, such models may not have relevance to the construction of a mental model of a domain (Kintsch, 1986). A psychological model of learning from texts, whether a single text or multiple texts, should include not only the text itself, but also the reader's previous knowledge and how the student uses that knowledge in constructing a new mental model.

Content and Disciplinary Knowledge

One goal of history instruction should be for the learner to construct a well-articulated mental model of history, understanding the interconnections between various events and actors. Taking the topic of the present study, the origins of the Vietnam War, a student should have an understanding of the relations between the U.S. election of 1964, U.S. views of communism during that era, Lyndon Johnson, the Viet Cong, and the Gulf of Tonkin Resolution. These understandings should be deep enough to consider how a possibly misunderstood incident, involving minor damage to two ships, could trigger a major conflagration. The mental model containing these understandings could be called "content knowledge" or knowledge about a particular domain (Stahl, Hynd, Glynn, & Carr, 1995).

Stahl et al., (1995) argue that, while content knowledge is important, it is not sufficient for the study of history. In addition, a person needs "disciplinary knowledge," or the ability to think like a historian, to evaluate materials and information in relation to their context and source, and to integrate this information into historical discourse (Greene, 1994). Wineburg (1991a, 1991b) gave eight historians and eight high school seniors a series of historical texts about the Battle of Lexington and had them complete a variety of activities, including "thinking-aloud" as they read, rating the trustworthiness of the documents, and evaluating the historical veracity of three paintings of the Battle. He noted that historians could be distinguished from students by their use of three processes:

- *Corroboration*, or comparing and contrasting documents with one another;
- *Sourcing*, or looking first at the source of the document before reading the text itself to consider how the bias of the source might have affected the content of the document; and
- *Contextualization*, or situating a text in a temporal and spatial context to consider how the time or place in which the document was written might have affected its content or the perspective taken.

The differences were not simply due to differences in content knowledge, since historians who did not know very much about the American Revolution still used the same

reasoning processes in their think-alouds; nor were the differences due to inability to detect bias. The college students in Perfetti, Britt, Rouet, Mason, and Georgi's (1993) study and the high school students in Stahl and Hynd's (1994) study were able to detect bias in sources.

Instead, the differences between the students and the historians seem to be tied to the ways in which historians and students view texts. Wineburg (1991a, 1991b) inferred that students tend to view texts as repositories for facts or as bearers of information, as they might well have, given years of exposure to a transmission model of learning. For example, they tended to rate textbooks as more trustworthy than source documents, a finding replicated by Perfetti et al. (1993) and Stahl and Hynd (1994). Historians tended to view texts as "speech acts," produced for a particular purpose by a particular person. To understand a text involves understanding both the person and the purpose, and to get at the "truth" hidden within the texts involves comparing various perspectives, with an understanding of who produced the various texts and why. Students in the Perfetti et al. (1993) study were able to grasp the basic story of the Panama Canal Treaty from documents describing the events leading up to the signing of the Treaty in 1903, but were less able to provide evidence about their stance on whether the treaty should have been signed.

Multiple Texts and History Learning

A number of educators have suggested that the single classroom text be supplemented with or supplanted by multiple original source materials (e.g., Perfetti et al., 1993; Spoehr & Spoehr, 1994; Wineburg, 1991a). Providing students with multiple perspectives on a particular event can aid them in constructing a richer and more detailed mental model of that event, thus enhancing content knowledge. Spiro, Coulson, Feltovich, and Anderson (1994) call the use of multiple perspectives "criss-crossing the landscape" and suggest that seeing an event through different perspectives is necessary to create a rich understanding of an event or concept. This use of original material forces students to construct links across information presented in different texts. This information and the links connecting the different sources are remembered better if students make their own constructions rather than rely on the constructions of a textbook author or teacher (Spoehr & Spoehr, 1994). The links based on this "criss-crossing" create a rich mental model, or what we are calling "content knowledge."

The use of multiple texts can also increase students' disciplinary knowledge. If we consider the tasks that Wineburg (1991a, 1991b) found to distinguish historians' thought processes and that of high school students—corroborating, sourcing, and contextualization—they only can be activated in students by providing opportunities to compare and contrast different source materials with different and independent viewpoints. The single, omniscient viewpoint of a textbook cannot easily be used to develop disciplinary knowledge, since there is nothing to which the student can compare the information. Thus, the student is usually unable to examine the bias of the textbook or the effects

of the time and place in which it was written, or to compare it to other sources. (However, McKeown, Beck, and Worthy, 1993, have developed procedures to elicit this information from critical examination of a single text.)

If conflicting information is presented in these texts, however, this may impede learning. Perry (1970) examined development of thought among male college students and found evidence for development ranging from a stance of looking for a single "right" answer to an understanding that knowledge is relative, depending on one's perspective, to the melding of information from different perspectives. Belenky, Clinchy, Goldberger, and Tarule (1986) replicated, with women, Perry's study. They found that stances of knowledge can move from (1) a belief that knowledge is received, or is transmitted from someone else, to (2) a subjective stance, in which knowledge is seen as subjective and relative, to (3) a procedural stance, in which rational processes are seen as a way to break through the subjectivity, to (4) a stance they call "constructed," in which knowledge is constructed through both rational processes and the acknowledgement of other perspectives. It is this last stance that we expect students to take when looking at multiple documents, but it is one that is typically achieved in the later college years or graduate school, after exposure to the more open-ended discussions typical of college classrooms. It may be unreasonable to expect high school students, who tend to be exposed to more lectures, to think like this, at least not without greater instruction in how to engage in this kind of thought.

Despite the theoretical sense that multiple sources can enhance learning, there is very little information on *how* readers synthesize information across texts. Spivey and King (1989) examined how sixth-, eighth-, and tenth-grade writers synthesized information from different encyclopedias. They found that older and more able students tended to be (1) more adept at using information that was repeated in all three texts and that was presumed to be more important, (2) better at reorganizing information from the different sources into a coherent whole, and (3) more aware of the reader's needs. Greene (1994) gave college juniors and seniors a task to either write a report or solve a problem in history. He found that students given the problem-based task were more likely to bring previous knowledge to their essays, to see the task as one of evaluation of the information in the articles, and to draw upon different kinds of information than the students who were given the report-writing task. The students who were asked to write a report had difficulty doing so, because they tended not to set their ideas in a context and justify the issues they chose to write about.

The purpose of this study is to examine the processes and outcomes of reading multiple original source materials. In this case, the materials relate to the Gulf of Tonkin incident and the resultant Tonkin Gulf Resolution passed by Congress, which eventually began the Vietnam War. We are specifically interested in the following questions: when given multiple historical source documents (1) can students develop a rich, mental model of a historical event? (2) what do students do with

the document information? (3) how do students integrate information across texts to form a coherent essay? and (4) do students engage in corroborating, sourcing, and contextualizing in evaluating historical materials?

Method

Participants

The participants were 19 students in two classes of tenth-grade Advanced Placement U.S. History taught by a single teacher. These students were enrolled in U.S. History with the expectation of being able to exempt a required course as part of their college curriculum. Therefore, only high-achieving students who were expecting to attend college were taking the class. The topics used in this study, The Tonkin Gulf incident and resultant Tonkin Gulf Resolution, were on the Advanced Placement exam, but students had not yet studied the incident and resolution in class. Researchers asked students a number of questions to determine their background. Approximately one-fourth of the students participating were African-American; the others were European-American.

This study was conducted in January, before the teacher began preparing classes for the document-based question on the Advanced Placement examination. Thus, these students had no direct instruction in how to integrate information across documents. Instead, the teacher used primarily a lecture mode, believing that "history is a story." The teacher was widely regarded as an excellent history teacher, with high percentages of students passing the Advanced Placement examination.

Procedures

Students who participated in the study met for 3 days in the computer room that was part of their school library. The librarian had equipped the room with 15 Macintosh SE30 desktop computers. As students came into the room on the first day of the study, the researchers handed each one a folder that included questionnaires, written directions for completing the study, and an introduction that assigned them to a topic and a purpose for reading. Researchers' distributed these folders in a stratified fashion to students upon entry, resulting in random assignment. Four conditions represented two purposes and two topics. Students were asked to read either (a) to form an opinion about the topic or (b) to be able to describe the topic. Students were also told that they would engage in a writing task related to their purpose for reading at the end of the study. Finally, we asked students to read texts about (a) the Tonkin Gulf Incident or (b) The Tonkin Gulf Resolution. For these topics, students could choose six and five texts, respectively.

Students filled out the background questionnaires, read the introduction that explained their task, wrote down everything they already knew about the topic they were assigned, completed the Gulf of Tonkin relationships task used as a pretest, and read the instructions for accessing the texts from the computer screen while the researcher explained those directions and answered questions. All students were familiar with the computers and with using the mouse, so they did not need basic directions for managing the computer. The researcher told students that they could read the texts in

any order they wished and take notes if they wished. After they had completed reading each text, they were asked to write a free recall (without looking back at the text they had just read), complete the Gulf of Tonkin relationships task, and fill out a questionnaire about the text. After completing those tasks, students could then proceed to their next chosen text.

Students started reading on the first day of the experiment, read through the 50-min period on the second day, and stopped reading on the third day, approximately 30 min before the end of the period. After students stopped reading the texts, they were told to read the directions for their writing task and to follow those directions. The directions asked students to *state their opinion* about either the Tonkin Gulf Incident or the Tonkin Gulf Resolution, or they asked students to *describe* the Tonkin Gulf Incident or the Tonkin Gulf Resolution. The students were allowed to consult their notes if they wished, but not allowed to return to the actual texts on the computer screen while they were writing.

Materials

Background questionnaire. The background questionnaire was the same as the one used in the pilot study during the previous year. The questionnaire asked students their political affiliation and their parents' political affiliation. It asked them whether they were liberal, conservative, or moderate on matters of national defense, the economy, and social issues. It also asked them about their stance on certain current affairs and issues debated in public forum, asked them to rate their knowledge of the Vietnam War, and asked them to describe their feelings about what was important to study in history. Finally, the questionnaire asked students to rate the United States Congress, United States newspapers, the President of the United States, army generals, historians, and history textbooks for their trustworthiness.

Prior knowledge writing task. After students filled out the background questionnaire and were assigned a topic, the following written directions were provided: "Please write down everything you know about your assigned topic. If you are not sure, then write down what you think you know." This task was scored for number of accurate knowledge statements and expressed as a percentage of accurate-to-total number of statements.

Gulf of Tonkin relationships task. In this task, students were asked to rate the strength of the relationship between all possible pairs of ten key words or phrases: The Gulf of Tonkin Resolution, The Gulf of Tonkin, North Vietnam, South Vietnam, United States Congress, President Johnson, Vietnam War, United States Forces, Defense, and Aggression. This task was given before any of the reading, as a pretest, and after each reading was completed, as a measure of growth as a result of that reading.

Students rated the pairs on a 1–6 scale with 1 being "not very related" and 6 being "strongly related." The purpose of this task was to determine the coherence (or harmony) of students' mental models before they read texts and afterwards, as a result of reading. Students were expected to have a more coherent way of rating the pairs after having read

texts. Of further interest was whether student responses would evidence steady growth and which texts were responsible for growth in coherence.

The measuring of "harmony" is described in Britton and Gulgoz (1991) and is expressed in the form of a decimal. For example, a harmony value of 1.00 would mean that an individual rated the relationships between pairs in such a way that there were no conflicts between ideas. A harmony value of .50, however, would mean that there was a moderate degree of contradiction in the way the pairs were rated. If a student rated the "Gulf of Tonkin Resolution" and the "President of the U.S." as strongly related, and "aggression" and the "President of the U.S." as strongly related, but "aggression" and the "Gulf of Tonkin Resolution" as not very related, the person's mental model of those three items would be considered inharmonious.

Texts. Students read multiple texts presented on Hypercard stacks on Macintosh computers.[1] Before reading any of the texts, students viewed a map showing Vietnam and the Tonkin Gulf, and read a 1½ card background information statement that described in objective terms the Tonkin Gulf Incident and resultant resolution. This background text provided an overview of the Vietnam War and the Gulf of Tonkin incident's role in that war. It was written to be neutral in terms of the two questions that were posed, providing facts that verified in all selections. The text is reproduced below:

> The war in Vietnam has been called the United States' longest war, because, even though the U.S. was not in combat the entire time, it was involved in the affairs of Vietnam for approximately 25 years, from 1950 to 1975. The U.S. became involved during the Truman administration, when it supported the French (who controlled the Vietnamese government) against a group of Communist rebels fighting for Vietnamese independence. By the time Lyndon Baines Johnson had taken over the presidency (in 1963), Vietnam had been divided into North and South, with the North being governed by the Communist president Ho Chi Minh, and the South being governed by a U.S.-supported president. In South Vietnam, a civil war had broken out in an attempt to topple the existing government and reunite North and South Vietnam under communism. This movement was led by a group the U.S. labeled the Viet Cong. The U.S. sent monetary aid, equipment, and advisors to the South Vietnamese government to support their fight against the Viet Cong and monitored, with concern, North Vietnamese support of the rebels. It was against this backdrop that the Tonkin Gulf incident took place.
>
> On August 2, 1964, shots were fired toward the *U.S.S. Maddox* by three PT boats while on patrol off the North Vietnamese coastline in The Gulf of Tonkin. Two days later, while the *U.S.S. Maddox* and a companion ship, the *U.S.S. C. Turner Joy*, were again on patrol, there were reports of another attack. President Johnson

[1]The stacks are available from the first author, 309 Aderhold Hall, The University of Georgia, Athens, GA 30602.

ordered a retaliatory strike and asked Congress to pass the Southeast Asia Resolution (also known as the Tonkin Gulf Resolution) to give him the authority to "take all necessary steps, including the use of armed force, to assist any member or protocol state of the Southeast Asia Collective Defense Treaty requesting assistance in defense of its freedom." This resolution was passed. Johnson used this approval to commit the U.S. to heavy involvement in the Vietnam War. "Hawks" (those who were supporters of the war) and "Doves" (those who were against the war) disagreed about what actually happened and about President Johnson's motivations in handling the incident.

After they read the background information, students were directed to look at a screen with two buttons, one directing them to documents concerning the incident and one directing them to documents concerning the resolution. They referred to their assignment sheets to see which question they were supposed to address. Clicking on a button led to a menu that presented the titles of their assigned readings. Students could browse the readings before deciding which ones to actually read. Because we wanted this task to be as natural as possible, we did not control the order of readings.

Six readings were about the Tonkin Gulf Incident and five were about the Tonkin Gulf Resolution. The topics were chosen because they have been hotly debated by historians and politicians. Different interpretations of the event and resolution exist, allowing text selection that represented several perspectives. It was the integration of various perspectives that the researchers wished to study. The texts chosen represented a blend of primary to tertiary sources that were as evenly distributed as possible in terms of their stances. The texts are listed on Table 1.

Since part of the focus of the study was to see which documents students would choose, texts representing a range of possible documents that might be used to study this incident were included. About half of the texts were judged to be pro-war, half anti-war. Histories (*Vietnam: A History* by Stanley Karnow and *The Pentagon Papers*), newspaper opinion papers, autobiographies of participants (Cmdr. James Stockdale and Dean Rusk), original documents (the text of the Tonkin Gulf Resolution and the telegram sent from the North Vietnamese protesting the earlier raids in the Gulf), and secondary sources were included. The intention was to ensure that all viewpoints were represented and that students had a choice of different genres and styles of documents. The information from a pilot study (Stahl & Hynd, 1994) was used in choosing texts that students rated as highly believable and those rated less believable.

As students read the texts, they had several options for help. For one, students could find out information about the author of the text. This information was basic and included the source of the document (newspaper, book, etc.) and the author's position (writer, former Army Colonel, Secretary of State, etc.). Furthermore, if the students put the cursor on selected vocabulary (mostly people and organizations), background information appeared on the screen. Students could also search for a keyword by choosing the *find* button and typing

Table 1. Texts Used in the Study

Title	Description
	Gulf of Tonkin Incident
Text of Telegram written by North Vietnamese to Protest the Mission of the U.S.S. Maddox	North Vietnamese call upon South Vietnamese to stop aggressive raids. Explain that U.S. ships were seen as aiding those raids.
Another Gulf, Another Blip On the Screen	An eyewitness account by James Stockdale, America's highest ranking Prisoner of War during the Vietnam War. He was flying over the Gulf of Tonkin and did not see a torpedo attack.
The Pentagon Papers	The official history of the event written shortly following the incident. It said that the Tonkin Gulf incidents were clearly aggressive acts on the part of the North Vietnamese.
Secrets of the Vietnam War	An excerpt from a privately published book written by a retired Army colonel, claiming that the North Vietnamese were primarily responsible for the incident and that the U.S. was not overly aggressive.
The Tonkin Gulf Crisis	An editorial analysis claiming that the Tonkin Gulf incidents were largely trumped-up by the U.S. as a way to widen the Vietnam Conflict.
Vietnam: A History by Stanley Karnow	An in-depth historical analysis that explains the events leading up to the U.S.'s interpretation of the Tonkin Gulf incidents that resulted in the Tonkin Gulf Resolution and ultimate widening of the Vietnam Conflict.
	Tonkin Gulf Resolution
The Tonkin Gulf Resolution	A copy of the actual resolution as voted on by Congress.
The Vote that Congress Can't Forget	A newspaper article that described members of Congress' retrospective thoughts about the Tonkin Gulf Resolution as they voted to allow the President to attack the Persian Gulf. Most of the congressmen said they regretted voting for the Resolution and that they didn't realize the effect it would have.
The Vietnam Hearings	Text taken from the Congressional Record describing Senator Fullbright's celebrated hearings where U.S. involvement in Vietnam was discredited. Dean Rusk, Senator Fullbright, and others were attempting to decide if spending more money on Vietnam was justified.
As I Saw It	An excerpt from Dean Rusk's autobiography which attempted to exonerate both the President and himself from accusations they they had acted hastily in their decision to escalate the war after the Tonkin Gulf incidents.
Vietnam War—A History	Same text as for Tonkin Gulf incidents.

in the word for which they were searching. They could take notes on the computer if they wished (although only three tried and all decided against it), and they could move freely backward and forward within and across texts.

Note-taking option. While students read each text, they could take notes on paper provided in their packet. Researchers and written directions explained to them that, although they were not required to take notes, students could use their notes for the final writing task but could not refer to the actual readings.

Evaluation sheet. Students were asked to answer these questions about each text: (1) What do you feel the author's purpose was in writing this?; (2) How useful would this be to help you learn about the origins of the Vietnam War? (rated from 1, "Not Very" to 6, "Very"); (3) How unbiased do you think this account is? (also rated from 1 to 6); (4) How difficult was this text to read? (1 to 6 rating); and (5) How interesting was this text? (1 to 6 rating). Students answered these questions before engaging in the free-recall task.

Free-recall task. This task directed students to: "Write down all the information you can remember from reading this text. Do not refer to your notes or the text before or during writing. Be as complete as possible." Students engaged in this activity after reading each of the texts.

Final writing task. Students were given a final writing task that mirrored their assigned purposes for reading. If students had been assigned to read in order to form an opinion about either the Tonkin Gulf incident or the Tonkin Gulf Resolution, they were asked to write about their opinions. If students had been assigned to read in order to describe the Tonkin Gulf incident or the Tonkin Gulf Resolution, they were asked to write a description. Students were given a 30-min period of time in class to complete this activity. All students finished before the 30-min time period was up.

Procedures for Analysis of Notes and Final Products

For the purpose of identifying processes that students used as they read each text and then formed an essay incorporating some or all of the texts they had read, a format was developed for recording the notes, text, and idea units from the essay so that their correspondences could easily be seen. Pages were divided into three columns: one for the text, one for the notes, and one for the essay. In the middle column, the notes were recorded (in idea units), in the order in which they were taken. In the left-hand column, the section of the corresponding text was recorded. Although judgment was sometimes needed to determine the textual basis for the notes, this task was relatively easy to perform because students generally took notes in the same linear order in which they read the text. Furthermore, the majority of their notes were paraphrases or copyings of the text. Idea units from the free-recalls were also recorded in this column, using the same procedures. However, the free-recalls were clearly marked as such so that they would not be analyzed as notes.

In the right-hand column, idea units from the final essay were recorded next to corresponding notes or text. Because the essay was

an incorporation of several different texts, idea units were recorded sometimes in several different places. If no corresponding note or text was found, the idea unit was placed at the end of the third column. Again, judgment was used in deciding whether or not an idea unit represented an idea taken from notes or text. In an attempt to be inclusive, if there was a possibility that students may have had a certain text in mind when they made the statement, it was placed accordingly. A portion of an analyzed set of notes and final product can be found in the Appendix.

After each student's notes and essay had been recorded in this manner, three researchers read all protocols. Using the method of constant-comparison (Glaser & Strauss, 1967), a system was created for categorizing idea units for the notes, free-recalls, and essays. To develop this system, three of the four researchers read through the protocols and discussed what they revealed about the students as they were taking notes and creating their final products. An attempt was made to codify these processes into a system that could be used reliably to categorize the processes found. A number of different systems were tried before an approach was found that could be applied with greater than 90% interrater reliability and that seemed to produce useful interpretations of the data. This categorization system is described below.

Notes and free-recalls. Each idea unit was classified as (a) copying; (b) paraphrasing; (c) reducing; (d) making a gist; (e) evaluating; or (f) distortion/misreading. An idea unit was classified as *copying* if it was word-for-word or nearly word-for-word, with close synonym replacement or minimal re-ordering. An example of copying is when the text said, ". . .Gulf of Tonkin Resolution passed by the Congress on 7 August 1964," and the notes said, ". . .*Gulf of Tonkin Resolution*—August 7, 1964 passed by Congress."

Paraphrasing was a more radical replacement of words that included within-sentence reduction or elaboration. An example of a paraphrase is when the text said, "Vietnamese coastal targets—this time the Rhon River Estuary and the Vinh Sonh radar installation, which were bombarded on the night of 3 August," and the notes said, "On the night of August 3, Vietnamese coastal targets were bombarded."

Reducing was described as a summarization process across two or more sentences, so that the writing contained markedly fewer words and details than the original. An example of reducing is when the text said, "At 1940 hours, 4 August 1964 (Tonkin Gulf time) while 'proceeding S.E. at best speed,' Task Group 72.1 (*Maddox* and *Turner Joy*) radioed 'RCVD INFO' indicating attack by PGM P-4 imminent," and later "Just before this, one of the PT boats launched a torpedo, which was later reported as seen passing about 300 feet off the port beam, from aft to forward, of the *C. Turner Joy*." The notes merely said, "On 4 Aug 1964, the *Maddox & Turner Joy* were attacked by PT boats, who launched a torpedo."

Gisting was described as radical reduction in which nouns were replaced with superordinates or more general terms. It was noted that gists were often blanket statements that were more topical in nature than reductions (such as

"the text was about the resolution"), or made blanket interpretations of details (such as "LBJ uses attack to get control of congress") when several paragraphs had described the President's dealings with Congress in getting the resolution passed.

Evaluating was described as stating an opinion about the ideas in the text that were not merely the copied opinion of authors or the opinion of people the authors described. For example, the statement "Johnson was an idiot" was classified as an evaluation.

Distortion/misreading was described as either inaccurate textual interpretations or statements that, although not evaluative, were simply not found in the text. An example of a misreading is when the notes said, "South Vietnam mistakes us for South Vietnamese ship," but the text said that the North Vietnamese mistook the *U.S.S. Maddox* for a South Vietnamese vessel.

As noted earlier, this coding system was developed after much discussion amongst three of the four researchers. After the system was developed, the researchers reached 92% agreement after coding notes on 5 of the 21 protocols. From that point, the researchers coded the remaining notes and free-recalls separately.

Final essay. The final essays were read and idea units coded as coming from a single text or two or more texts. If a significant number of statements came from two or more texts, it was assumed that students were either integrating ideas across texts or paying attention to information that was repeated across texts. Each idea unit was coded also as a copy, a paraphrase, a reduction, a gist, an evaluation, or a misreading, as was done with the notes and free-recalls. The purpose of this categorization process was to analyze what processes students were using to form a coherent essay.

In addition, the order of statements in relation to the order of the texts the students read was noted and Kendall's Coefficient of Concordance (W) to obtain a measure of the overall agreement in order was performed. This coefficient was used to determine whether students were radically restructuring ideas or merely reporting them in the form in which they were first perceived. A low W indicated that students were reordering from the texts in the final product. A high W indicated that they were generally preserving information in the order that they had read it. Finally, a ratio of information found in the text to that which could not be found in the text was calculated. This ratio revealed whether students were sticking to the task of describing or stating their opinions, as they were assigned. It also revealed whether students who were asked to state opinions would back up these opinions with factual information.

Sourcing, corroboration, and contextualization. Wineburg (1991a, 1991b) observed that, when thinking about information in the texts they read, historians used sourcing (looking first at the source of the document before reading the text itself to consider how the bias of the source might have affected the content of the document), corroboration (comparing and contrasting documents with one another), and contextualization (situating a text in a temporal and spatial context to consider how the time or place in which the document was written might have affected its content or the perspective taken).

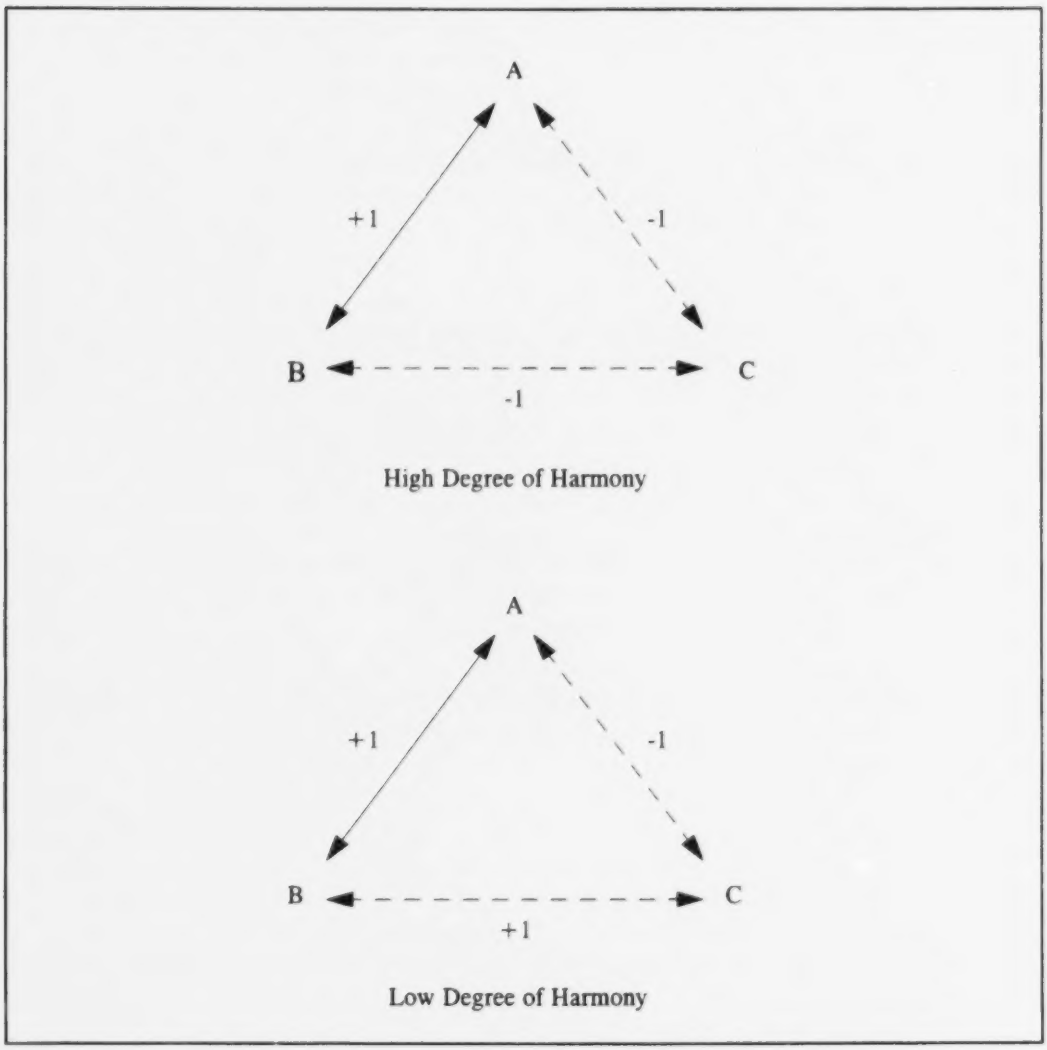

Figure 1. Model of Harmony

By contrast, students used these methods to a lesser degree, if at all. In this study, the researchers looked for instances of sourcing, corroboration, and contextualization in the notes, in the free-recalls, and in the final essays.

Table 2. Harmony Ratings

	Harmony levels		
	Read 1 Text	Read 2 Texts	Read 3 Texts
Pretest	.67	.67	.71
Text 1	.70		
Text 2		.76*	
Text 3			.81*
	Harmony after reading specific texts		
Pretest	.67		
Pentagon Papers	.72		
Vietnam Hearings	.71		
Tonkin Gulf Crisis	.75		
Another Gulf, . . .	.78		
Text of Telegram	.71		
Vietnam: A History	.81*		
Vietnam Hearings	.77		
As I Saw It	.79		
The Vote that Congress Can't Forget	.75		
The Gulf of Tonkin Resolution	.74		

*$p < .05$

Results and Discussion

Can Students Develop a Rich, Mental Model of a Historical Event?

The data from the relationships task were used to track how students developed a mental model of the events surrounding the Gulf of Tonkin incident and the Gulf of Tonkin Resolution. Two approaches were taken to examine the development of a mental model. The first approach examined the growth of harmony, which researchers used as a proxy for the internal consistency of the mental model developed by the students. The second approach compared the structures generated by the students to those of experts, to trace the growth of students' mental models toward those held by experts.

Harmony. As students learn, they sort out internal contradictions between different ideas and begin to generate stable relationships between ideas. At the top of Figure 1, concepts A and B are strongly related, but A and C are

negatively related. Since B and C are not related, this model shows a high degree of harmony. In the bottom part of Figure 1, B and C are strongly related—a relation inconsistent with the others—indicating a low degree of harmony. As noted earlier, if a student feels that "South Vietnam" and "Congress" are both strongly associated with "Defense," then they should be closely associated with each other. This would indicate a high degree of harmony. If the concepts are seen as distantly related, this would indicate a low degree of harmony. Theoretically, the students with the best developed mental models would have high internal consistency or harmony.

The harmony ratings (using Kintsch's 1994 system of calculating those ratings) are shown in Table 2. As noted in Table 2, there was a significant growth in harmony from the pretest to after the second reading and after the third reading. The growth in harmony after the first reading was not statistically reliable. Comparing the growth in harmony after each reading, only the difference in harmony between the first and second reading was statistically significant. There was a small, further increase in harmony after a third reading. However, the number of students who got to the third reading (17) was small, but the absolute difference was low as well (.3733 vs. .3786). This suggests that a student needs to read at least two different texts to develop a coherent mental model and that the majority of growth occurs after two readings.

To examine the effects of individual texts on the growth of harmony, the gains in harmony from the students' prereading ratings after they read each of the texts were examined. Of the 10 texts used, only the section from the history text, *Vietnam: A History*, produced a significant gain in harmony by itself. This might be expected, since it was the longest and most detailed text we used.

Expert ratings. Another way to examine mental models is to compare the structures generated by the students with the structures generated by experts. This suggests that *knowledge* consists of knowledge of relations among concepts and that, as a person's knowledge grows, his/her knowledge of the relations among concepts will resemble that of experts. Three expert raters were used to generate structures, using the same terms and procedures that were used with the students. The first rater was the students' high school teacher, an experienced history teacher. The second rater was an amateur military history "buff" who read extensively about the War in Vietnam. The third rater was one of the authors of this study, who read the documents thoroughly and responded to the task based on her reading of the texts. These raters rather than studied "experts" on the Vietnam War were chosen because they represented the level of expertise expected from the students. There was not enough information in the texts to allow the students to obtain as full a representation of the events in Vietnam as a scholar would. Because these tasks focused on a small incident embedded within a larger context, it would be unrealistic to compare the knowledge obtained from these readings to that of scholars who have been immersed in the larger context. The level of expertise our raters had was about that which could be reasonably expected on this task.

Table 3. Correlations with Expert Raters

	Correlation with Average of Three Experts	*r* with Expert 1	*r* with Expert 2	*r* with Expert 3
Pretest	.26	.21	.15	.29
After Text 1	.42*	.33*	.26*	.41*
After Text 2	.36*	.28	.23*	.37*
After Text 3	.43*	.33*	.30*	.41*
After Text 4	.38	.33*	.23	.35

*Correlation significantly different from pretest $p < .05$.

All three experts tended to cluster the terms around two axes, one separating the terms "Aggression" and "Defense," and the other roughly separating terms into domestic (U.S. Congress, U.S. Forces, etc.) and foreign (N. Vietnam, S. Vietnam, Gulf of Tonkin, etc.). Experts tended to have a strong separation between "Aggression" and "Defense," and clustered the other terms in the middle, roughly equidistant between these two poles.

In this model, a gain in knowledge would be evidenced by an increase in the students' correlation of their mental structures with that of the experts. Examining these correlations, shown in Table 3, there was a significant growth in knowledge after the first reading, but no significant gain subsequently. The initial correlations between the students and the individual experts ranged from .15 to .29, and initial correlation between the students and the composite was .26. These are small and not statistically reliable, suggesting that the students' initial knowledge was low and essentially random. The gain to .42, a moderate correlation, suggests that students learned some of the initial relationships after a single reading. Since subsequent readings tended to present the same facts from different perspectives, it is not surprising that there was little gain after these readings. There is also some evidence that students read the first reading more closely than subsequent readings.

In contrast with the experts, students tended to cluster "N. Vietnam," "Gulf of Tonkin," and the "Gulf of Tonkin Resolution" with "Aggression," and "S. Vietnam," "U.S. Force," and "Congress" with "Defense." This may reflect a different world view than that of our experts, who all lived through the Vietnam era. These students tend to see the United States and its allies in a positive light, and its enemies in a negative light. By contrast, the experts tended to view both sides with a more balanced regard, as neither side being more defensive or aggressive than the other.

One explanation for the lack of growth after a first reading may be in the nature of the texts and the task. We deliberately chose texts that

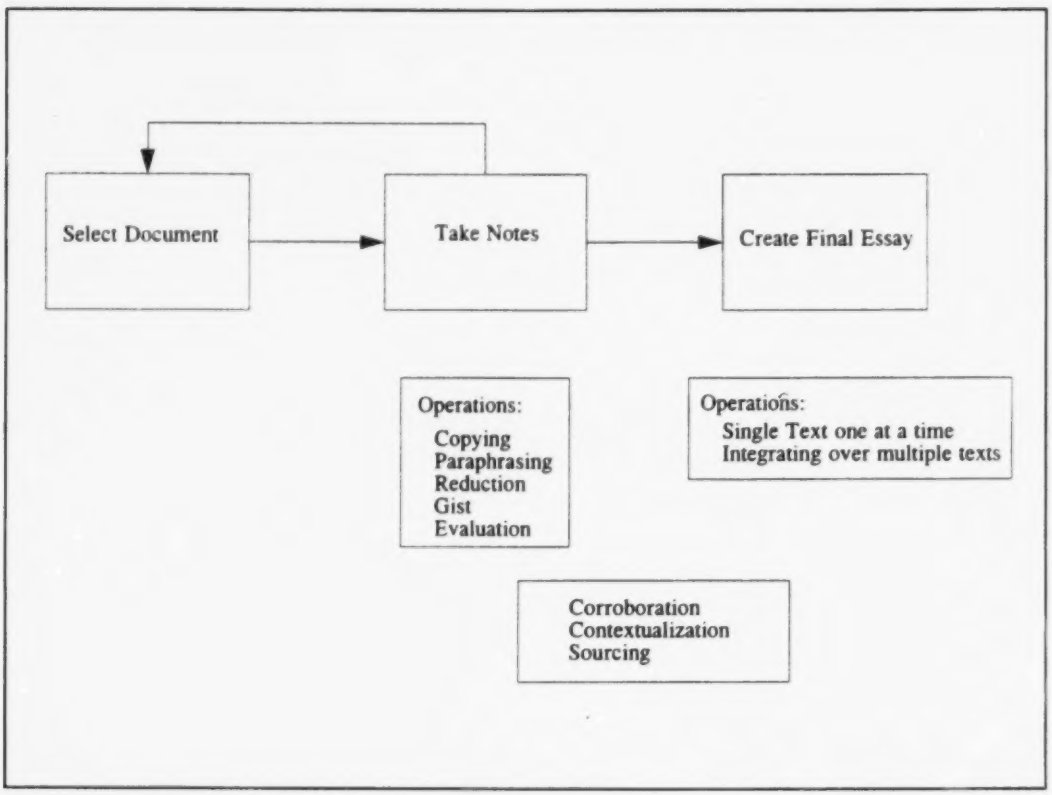

Figure 2. Model for Analysis of Notes and Final Product

contradict each other. It could be that students would read a first text to get the basic facts, and then go through each subsequent text, filtering out the contradictions and looking for overlap. When reading contradictory texts, students may be "averaging" out the opinions, trying to stay with opinions agreed-upon by more than one author, rather than constructing an increasingly complex mental model that might be closer to that of the experts. This needs to be tested in a more comprehensive study.

Background information. The students were given an extensive background questionnaire, asking them for information such as their political orientation, their parents' political orientation, their views about current events, and their views of the reliability of various people and institutions such as Congress, the President, historians, and so on. No relation was found between any of the background variables and students' responses to these measures or any of the other measures.

As discussed below, students with moderate to high knowledge of the Vietnam War did take different types of notes than other students.

What Do Students Do With the Document Information?

As noted above, an attempt was made to follow the flow of ideas from each document, through the notes, to inclusion in the final product. The basic model hypothesized is in Figure 2. This model suggests that students initially select ideas from the text as they are reading, deciding which ideas are important and which are not. They may make a note of a selected idea, either copying it, paraphrasing it, reducing two or three sentences, or reducing a paragraph or more into a single gist statement. They may also note an opinion or reaction to information in the text. In producing the final product, they use similar operations, with ideas from a single text or ideas combined or repeated from multiple texts.

Choosing texts. The excerpts from two texts—*Vietnam: A History* and *Secrets of the Vietnam War*—were chosen by more students to be read first than any other texts. Each was chosen by about one-third of the subjects; the remaining texts were chosen by the remaining third of the subjects. It is speculated that the history text was chosen because it seemed to provide an overview, and because students would perceive it as neutral in tone. It is not known why the "Secrets" text was so popular. It seemed to be an important source of information in the students' final products as well.

Selecting information. When reading a text, students must first select which information is important. Given that these were "natural" texts, not especially created for this study, they varied considerably in how well they were constructed. Two texts were especially poorly constructed. *The Pentagon Papers*, for example, is a detailed history of the Vietnam War, written for internal purposes by the Army, and contains many gaps (indicated by a notation reading "Several Paragraphs Missing"). The text is written for bureaucrats and is highly inconsiderate of the reader (Anderson & Armbruster, 1984). *The Senate Hearings* are a transcript of the hearings, written in a play format. Other texts were written for purposes other than those given to our participants. Commander Stockdale's account used his experiences in the Tonkin Gulf to comment on the unreliability of radar data in a more recent incident. The difficulties of using naturally occurring documents is that the student has to sift through a great deal of irrelevant information to find what is important.

There were differences in how consistent students were in selecting information among documents. A tally was made of how many students annotated each statement in each of the texts to investigate selection patterns. In some documents, students tended to select the same idea units in their notes. These tended to be shorter and more focused documents. The statements themselves tended to be clear, strongly stating an opinion about the incident or the resolution. For example, in *The Tonkin Gulf Crisis*, 7 of the 9 students who read it annotated the statement, "The accumulated evidence makes it reasonably certain that the alleged North Vietnamese PT boat attack of Aug. 4 was a figment of the U.S. government's

imagination." Two other statements were annotated by 5 students and another by 4.

In other documents, students diverged widely in terms of what information they selected. In *The Pentagon Papers*, one statement was annotated by 4 of the 6 students who read it. ("Upon first report of the PT boats' apparently hostile intent, F-8E aircraft were launched from the aircraft carrier Ticonderoga, many miles to the south, with instructions to provide air cover but not to fire unless they or the *Maddox* were fired upon.") A total of 53 statements were annotated, but no other statement was noted by more than half of those reading. Few students read the *Vietnam Hearings*, the other text we judged to be poorly structured.

Thus, it appears that the nature of the text affects how students select information. Students tended to be more consistent in what information they selected from short, well-constructed texts. In these texts, they tended to choose strong, clear statements of position. In *The Pentagon Papers* excerpt, a longer, less well-structured text, students chose many different statements, with only one statement chosen by more than half of those who read it.

However, students rarely chose irrelevant information. In the Stockdale article, no student mentioned the current incident, annotating only information dealing with the Gulf of Tonkin incident. In *The Vote that Congress Can't Forget*, which looked back on the Gulf of Tonkin Resolution by contrasting it with the authorization of the Persian Gulf War, only 2 students annotated information dealing with the Persian Gulf War. Thus, students were good at filtering out information they did not need.

Does the Task the Students Are Given Influence Their Processing?

It was hypothesized that the type of task, either describing or forming an opinion, would affect processing, as evidenced by the notes the students took. Students given a task of describing would concentrate more on details in their notes and might include more copying and paraphrasing. Students who were asked to form an opinion might reduce larger chunks of text into main idea statements and might include more statements classified as "reduction" or "gist" in their notes.

Of our 19 students, only 17 took notes which could be analyzed. Of those 17, 9 were asked for an opinion and 8 were asked to describe either the incident or the resolution. Differences were examined using Discriminant Analysis, a fairly sensitive multivariate analysis technique. Neither this analysis nor other appropriate analyses found significant differences between those given different tasks and those given different topics.

The lack of differences is surprising, because students who were asked to form an opinion were expected to concentrate on more global information and construct more gist statements and evaluative statements, and students who were asked to describe the incident to concentrate on details and copy more information directly or in paraphrase. Even those asked for an opinion included few evaluative statements. Of the 8 students asked for an opinion, there were only 11 evaluative statements made in the notes; an additional 5 evaluative statements were made by the 9 students asked to describe. This is

a very low number of the hundreds of notes made by the students.

The means, however, mask strong individual variations in how students approached the task. Some students took copious, detailed notes, no matter which task they were assigned. As noted on Table 4, two students who were asked to write a description (#19 and #28) consistently took many notes, as did #18, who was asked to form an opinion. Others tended to write gist statements, condensing a great deal of information into brief, even telegraphic, notes, such as #42 (opinion) and #33 (description).

One student (#43) made considerably more evaluation comments than the others. This student indicated that he had relatively high knowledge of the Vietnam War on the preassessment, and many of his comments reflect that knowledge. For example, his evaluative comments tended to reflect a strong bias such as "[The] U.S. was not wrong in firing on the Vietnamese," "Vietnamese started War," "Johnson's an idiot," and "So it sounds like its a bunch of idiots playing with their guns." This bias appeared to be based on foreknowledge, rather than developed through reading. This student also read through all six texts in the time allotted for the study, the only student to read this many.

To investigate how the level of prior knowledge affected the processing of these documents, notations were made of each student's rating of knowledge of the Vietnam War and number of correct idea units in the original assessment of knowledge about the Tonkin Gulf incident or Resolution. No consistent pattern was found, however, possibly because very few students knew anything about the Gulf of Tonkin. For example, the average number of correct idea units about the Gulf of Tonkin was less than one. One would expect that students with high levels of knowledge would take more gist-like notes. However, this sample did not include enough high-knowledge students to observe that phenomenon.

The order of the texts also seemed to affect how many notes were taken. A repeated measures Analysis of Variance (ANOVA), looking only at the first three text readings (since few students read more than three texts), found a significant difference among readings [$F(2, 32) = 9.07$, $p < .001$]. Students averaged taking 11 notes for the first text and 5 apiece for the second and third texts read. The greater amount of notes taken for the first text may indicate that more effort was expended in reading the first text. Recall that only after the first text was read did students make significant growth toward the experts' knowledge structure. There appeared to be little effect of task or topic on readings of the different texts.

How Do Students Integrate Information Across Texts to Form a Coherent Essay?

Free recall. Students were more likely to reduce and make gist statements in the free-recalls than in the notes, regardless of whether they were asked to write a description or form an opinion. This behavior seems reasonable in that students were relying on memory and were not able to paraphrase easily information in the texts. It also argues for the idea that students processed information in similar ways, regardless of the final task.

Table 4. Processes Used in Notes (First Three Texts Read)

Subject ID	*Task*	*Topic*	*Text Read*	*Copying*	*Paraphrasing*	*Reduction*	*Gist*	*Evaluation*
First Text Read								
28	Description	Incident	Karnow	14	8	2	1	1
34	Description	Incident	Karnow	0	1	0	0	0
12	Description	Incident	Pentagon Papers	7	10	3	2	0
29	Description	Incident	Karnow	1	3	0	9	0
27	Description	Resolution	T.G. Crisis	1	8	0	1	0
19	Description	Resolution	Rusk	9	4	1	2	0
33	Description	Resolution	Resolution	0	2	0	1	0
23	Description	Resolution	The Vote	0	1	0	2	0
37	Description	Resolution	Rusk	2	9	2	2	0
Description Means				3.78	5.11	0.89	2.22	0.11
42	Opinion	Incident	Karnow	1	3	6	10	3
41	Opinion	Incident	Secrets	0	0	0	0	0
18	Opinion	Incident	T.G. Crisis	9	6	0	0	0
26	Opinion	Incident	Papers	7	3	1	0	1
16	Opinion	Incident	Secrets	7	3	1	0	1
6	Opinion	Incident	Papers	7	5	1	0	0
40	Opinion	Resolution	Secrets	0	5	0	3	1
30	Opinion	Resolution	Resolution	0	9	1	0	0
Opinion Means				3.00	3.88	1.13	1.63	0.63
Second Text Read								
28	Description	Incident	Telegram	1	5	1	0	0
34	Description	Incident	Stockdale	4	2	1	2	0
12	Description	Incident	Secrets	5	0	2	0	0
29	Description	Incident	Secrets	0	4	6	0	0
27	Description	Resolution	Resolution	3	0	0	0	0
19	Description	Resolution	Resolution	2	2	0	0	0
33	Description	Resolution	Rusk	0	4	1	3	0
23	Description	Resolution	Resolution	1	1	0	0	0
37	Description	Resolution	?	0	0	0	0	0
Description Means				2.00	2.25	1.38	0.63	0.00

Table 4. (Cont.)

Subject ID	*Task*	*Topic*	*Text Read*	*Copying*	*Paraphrasing*	*Reduction*	*Gist*	*Evaluation*
Second Text Read (cont.)								
42	Opinion	Incident	Secrets	1	4	2	0	3
41	Opinion	Incident	Telegram	0	0	0	0	0
18	Opinion	Incident	Telegram	1	6	1	0	0
26	Opinion	Incident	Stockdale	0	0	0	3	1
16	Opinion	Incident	T.G. Crisis	5	0	1	0	1
6	Opinion	Incident	Telegram	0	4	0	0	0
40	Opinion	Resolution	Papers	0	3	0	1	1
30	Opinion	Resolution	T.G. Crisis	1	4	0	0	0
Opinion Means				1.00	2.63	0.50	0.50	0.00
Third Text Read								
28	Description	Incident	T.G. Crisis	4	5	0	0	0
34	Description	Incident	T.G. Crisis	5	1	1	0	0
12	Description	Incident	T.G. Crisis	2	0	0	4	0
29	Description	Incident	T.G. Crisis	0	1	0	4	1
27	Description	Resolution	Rusk	3	4	0	1	0
19	Description	Resolution	Karnow	6	4	1	0	0
33	Description	Resolution	?	0	0	0	0	0
23	Description	Resolution	Rusk	0	0	0	0	0
37	Description	Resolution	?	0	0	0	0	0
Description Means				3.33	2.50	0.33	1.50	0.17
42	Opinion	Incident	Papers	2	5	3	3	4
41	Opinion	Incident	Stockdale	0	0	0	0	0
18	Opinion	Incident	Stockdale	1	5	2	0	0
26	Opinion	Incident	Telegram	0	0	3	0	0
16	Opinion	Incident	Papers	2	3	1	0	0
6	Opinion	Incident	T.G. Crisis	0	0	0	0	0
40	Opinion	Resolution	T.G. Crisis	1	2	1	1	0
30	Opinion	Resolution	Rusk	2	8	2	1	0
Opinion Means				1.14	3.29	1.29	1.14	0.57

Table 5. Processes Used in Final Product

	Single Text			Multiple Texts		
	Paraphrasing	Reduction	Gist	Paraphrasing	Reduction	Gist
Opinion	.07 (.07)	.00 (.00)	.03 (.07)	.01 (.04)	.05 (.08)	.60 (.24)
Description	.20 (.16)	.13 (.13)	.22 (.27)	.10 (.13)	.07 (.08)	.23 (.22)

N.B. Numbers are mean percentages. Standard Deviations in parentheses

Final product. On the final product, students tended to stick to the task. As can be seen in Table 5, the students who were asked to describe engaged more in paraphrasing, reducing, and making overarching gist statements about a particular text than did students who were asked to form an opinion. Students who were asked to form an opinion rarely paraphrased or reduced. Rather, their final essays were replete with evaluative/gist statements such as, "I believe that the U.S. was too quick to pass the Tonkin Gulf Resolution." These statements can only be seen as conclusions reached from reading more than one text, although evidence backing up these statements was scanty at best. Note the low number of paraphrased statements or reductions relating to either one text or a combination of texts. These types of statements would count as evidence supporting their opinions. Interestingly, student #33 (mentioned previously), who wrote many gist statements in his notetaking despite the fact that he was asked to write a description, stuck to description when he composed his essay. Similarly, many students who mostly copied or paraphrased their notes, despite the fact that they were asked to write an opinion, wrote opinion-like statements when they composed their essays.

Students who were asked to write an opinion tended to move away from the text, toward broader generalities and statements without any apparent grounding in the texts read. Even though they indicated a depth of reading through their notes, their final products seemed to disregard that depth. For example, student #6 wrote:

> (1) *My opinion is that the* USS Maddox *did get attacked by the North Vietnamese the first time* (2) *but was not attacked in the second "incident"*. (3) The reason for the first attack was that the North Vietnamese thought the *Maddox* was a South Vietnamese ship (4) and since the South had attacked the night before they defended themselves. (5) Later on the South Vietnamese attacked the North Vietnamese again. (6) The *Maddox* was again patrolling (7) and the US government thought prematurely that the North Vietnamese would once again attack. (8) *The US government reacted*. (9) *I'm not sure if Johnson lied or what happened*. (10) *In my opinion, something wrong happened. It*

> *sounds like it might have bene [sic] the US fault.* (11) *It might be this because several of the texts said the same thing.* (12) That nothing was out there when the *Maddox* and the *Turner Joy* were patrolling. (13) *I am not sure exactly why the USA would do this.* (14) *They might not have.* (italics and numbering added)

This student took notes throughout the text, but half of his statements (italicized) could not be reconciled directly with texts that he read. He seemed to view the task of giving his opinion as being dissociated from getting evidence from the text to support that opinion. The first statement is a clear thesis statement; the following statements do support that thesis. However by the eighth statement ("The US government reacted"), he becomes vague and speaks in generalities, ending in confusion. This may be because he lacks experience with writing coherent texts using an argument structure or he is still confused by the contradictory texts and has not yet examined the evidence to form an opinion.

As might be expected, those who wrote descriptions tended to stay closer to the readings. These students provided few evaluative statements. An example would be that of student #28.

> (1) The Tonkin gulf incident occurred due to a series of events such as the first battle in which the *Maddox* was legitimately involved in (due to the attack made by the North Vietnamese). (2) The second battle which some feel never really happened because no one actually saw any PT boats, also had a large effect on the Tonkin gulf situation. (3) It led to Congress passing of the Tonkin gulf Resolution, the retaliatory acts wanted by the Sec. of Def. and other officials that were allowed by President Johnson. (4) These things combined led to the N. Vietnamese feeling that war would occur in the South and moved troops down the Ho Chi Minh trail, resulting in what would possibly be interpreted by US officials as aggression. (Numbers added)

Statements 2 and 3 were supported by three references in the text apiece; the fourth statement by a section in "The Tonkin Gulf Crisis," the last text read. This student took copious notes; however, very few of the 42 annotations were actually used in this short essay.

Integration. Students did appear to use more than one source of information in forming their final essays, and they engaged in rearranging ideas from single texts as they wrote. To examine how students integrated information across texts, two analyses were conducted. First, each statement in the text was categorized as to whether it had one source in the readings or whether the idea could be found in multiple readings. (This may have overestimated the number of statements classified as coming from multiple readings, since an idea was categorized as coming from multiple readings whether or not that idea appeared in the student's notes in two places.) Students who were asked to write an opinion tended to use more ideas that came from multiple texts (64% of statements) than students who were asked to write a description (40% of statements). Students who were asked to write descriptions used

more ideas that could only be found in a single text (55% of statements) than students asked for opinions (10% of statements).[2]

Next to be considered was the ordering of ideas in the texts read and in the final product, using only those ideas that could be identified with a single text. The order of the statements in the final products was compared with the order of those statements in the texts the students read, in the order in which they read them. Kendall's Coefficient of Concordance (*W*), a measure of interrater reliability, was used to compare the different orderings. If students had merely written ideas from the texts in the order in which they were presented, the mean Coefficient of Concordance would have been 1.00. However, the mean of the total final essays was .76, and the range was between .38 and 1.00. Most of the participants made little or moderate change in how the texts were used in the essay compared to how they were read. Only 2 of the 19 students made drastic reorganizations. There was essentially no difference in the coefficients of the students who were asked to write descriptions and those who were asked to write opinions. The essays from both groups were coherent in that they had discernable beginnings, middles, and endings.

These results argue for the idea that students reading multiple texts are able to form more elaborate networks of ideas, in that they seem to integrate information across multiple sources. For example, Subject #25 (Opinion on Resolution) began his essay with a thesis statement, *"President Johnson was definitely justified when he asked Congress to pass the Resolution."* His next sentence, *"He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies,"* seemed to draw from two sections of the Resolution as well as two sections of Dean Rusk's autobiographical recollections. The remainder was drawn also from these two sources. Although he read and took extensive notes from a third source—a newspaper opinion piece about the Resolution—he did not take any ideas from that source for his final product. This opinion piece argued against the position the student was assigned to take, and thus was ignored in the final essay.

Do Students Engage in Corroborating, Sourcing, and Contextualizing in Evaluating Historical Materials?

As noted on Table 5, which contains the number of comments in the notes classified as either sourcing, corroboration, or contextualization, few students had comments which could be classified as reflecting the processes used by the historians studied by Wineburg (1991a). What is interesting is that the two students who included a great many gist statements also tended to include some sourcing statements. Student #42, who included 12 gist statements, also included 10 statements dealing with the source of the documents. This student wrote a lot of notes, covering four texts. His notes tended to be telegraphic, just a few words to cover the main ideas. He began

[2]Totals do not add up to 100% due to rounding and framing and evaluative statements, which would not be classified here.

staying closer to the text, paraphrasing mainly in the first two texts he read. Toward the end, though, he produced global statements about the text, which were classified as gist statements, such as "This text was basically dialogue that outlined how the Senate felt about the current situation in the Gulf of Tonkin" (notes on *The Vietnam Hearings* excerpt). This was also classified as "sourcing," since it makes reference to the text, but this is not sourcing in the same sense that Wineburg suggests. Instead, the student refers to the text and the participants, not from foreknowledge of their roles, but as placeholders representing sides.

That students included very few comments which could be considered as sourcing, corroboration, or contextualization suggests that they lack the knowledge of the discourse patterns of historical analysis. As Wineburg (1991a) points out, professional historians approach the task of reading documents as members of this discourse community, but high school students do not. This was true even for historians who scored lower than high school students on a test of the factual content. This knowledge of discourse patterns represents the disciplinary knowledge of history, or the ability to think as a historian might, and may need to be taught directly.

Conclusion

The focus of this research was to study the processing of information by students who read multiple historical documents about a controversial event in history—the Tonkin Gulf incident and the subsequent Tonkin Gulf Resolution that led to heavy U.S. involvement in the Vietnam Conflict. One aspect of the study was to understand what happens to students' mental structures when they read more than one text about an incident, particularly when those texts propose alternative interpretations of the event. Also of interest was whether students would employ different strategies for processing the texts if they were given different purposes for reading.

This study was intended to be exploratory. There were a large number of possible variations in the study—students differed in terms of task and topic, and also differed in what texts they read and in which order. Because of the large number of variations, it was difficult to make definitive statements. Instead, a possible model of students' processing of multiple texts is proposed, based on interpretation of the data, followed by a discussion of why studying documents alone might not lead to the disciplinary knowledge proposed by Wineburg (1991a) and others.

The proposed basic model suggests that the process can be broken down into *selection* of ideas in each text read, *processing* of ideas within that text, *constructing* a mental model of the information, and *integrating* ideas across texts to produce a final product. These will be discussed in turn.

Selection

The students in this study tended to be influenced strongly by text features in their selection of ideas. Students consistently chose to note the same ideas from short, well-structured texts. These ideas tended to be

clear, strong statements of opinion or topic sentences encompassing a great deal of detail. With long and badly structured texts, such as *The Pentagon Papers*, few students chose the same ideas. With such texts, it was difficult to pick main points, since there was so much detail and little attempt at organizing it. Original source documents, however, tend to be more like *The Pentagon Papers* and *The Vietnam Hearings*, than shorter, more focused pieces. Students need to learn how to cull information from longer documents if they are going to be used in units such as this.

However, students were able to concentrate on relevant information. In two pieces written for a purpose other than the assigned purpose, students consistently ignored irrelevant information, focusing instead on information suited to the purpose.

Processing

The kind of task had little effect on how students read the information in the different texts. Students who were asked for an opinion did not differ from those who were asked for a description in the types of notes they took. There were few evaluative statements given, even by those asked to write an opinion. Students tended to take many more notes on the first text than on subsequent texts, suggesting that they were expending more cognitive effort in constructing a mental model of the information in the first text.

There were, however, strong individual differences in notetaking. Some students tended to take copious, detailed notes, relying on copying and paraphrasing. Others tended to rely on gist statements, noting only main points, often only telegraphically. These differences in notetaking strategy do not seem to be related to the task or to which text was read, but seem to be an individual difference.

Constructing

Analysis of the students' ratings of relatedness among key terms suggests that students' mental structures tend to grow in two ways while reading multiple texts. Students' structures tend to be more internally consistent after reading a single text, and then still more consistent after reading a second text. The history text tended to produce the greatest growth in harmony or internal consistency. Students' structures also tended to become more similar to those of experts after reading a single text, with no further growth after reading two or more texts.

Similarity to experts was used as the measure of growth of knowledge. The results suggest that students did not grow in their knowledge after more than one reading, but they did become more consistent in their understandings. This lack of growth may be because they simply did not process the subsequent texts as well as the first. There is evidence of a clear decline in notetaking after the first text read. Another complementary explanation lies in the nature of the texts read. Because the texts were chosen to contradict each other, students may have looked for overlap between texts, rather than for new knowledge. The overlap would reinforce the basic knowledge acquired by reading the first text, but might not add very much to the student's understanding.

Integrating

The type of task students were given strongly influenced their final product. Students who were asked for a description tended to stay close to the texts, with most of their statements coming clearly from information provided, usually in a single text. Students who were asked for an opinion tended to produce more global statements not clearly tied to any single text, but that could be found either in multiple texts or not in any text.

Limitations

This was intended as an exploratory study. Further work to examine the model discussed above is needed. First, systematic variation of the texts that students read should be employed. The present study investigated what would happen if students were given freedom to choose whatever texts they wanted, and if there was a pattern in their choice. Variation of the texts in a principled manner would explore how different types of texts—histories, opinions, source information—affect students' learning. Second, background knowledge of the topic should be varied. Students' knowledge in the present sample was uniformly low. A great deal of effort seemed to be expended on constructing a basic understanding of what went on in the Gulf of Tonkin and in the U.S. Senate during the discussion of the Resolution. This may have hampered students in evaluating the information in the texts. Third, the problem of the Tonkin Gulf is a problem of perception: which of two clearly contradictory sides is correct? The processes described here might not be found in a less polarized topic, such as the Panama Canal Treaty as studied by Perfetti and his colleagues (1993). Comparing different types of historical problems is needed to examine their separate effects on students' learning.

Thinking Like a Historian Using Multiple Source Documents?

Some students did engage in some of the processes described by Wineburg (1991a) as being typical of professional historians: contextualizing, corroborating, and sourcing. Engaging in these processes was also an individual difference: some students did so frequently, but most did not evidence these processes at all.

For most of these students, presentation of multiple texts did not encourage them to think like a historian. In fact, the greatest growth of knowledge came after reading the first text, and the text that had the greatest influence on growth of harmony was a well-organized history textbook, albeit a text devoted entirely to Vietnam. Students read the first text to get basic facts and information, and read subsequent contradictory texts in an attempt to sort out that information.

One reason many students did not seem to develop disciplinary knowledge from reading multiple texts was their lack of initial knowledge about the topic. Students' initial reliance on the history text and their tendency to take paraphrase-type notes may have been reflections of their need to gain a literal understanding of the content before attempting to produce an opinion. Alexander and Judy (1988) argue that students become able to use more sophisti-

cated strategies for learning new information when they already have some content knowledge. The students we studied may have been taking notes in paraphrase fashion initially because they lacked background knowledge and were reading to gain this knowledge, regardless of the final task they had been assigned. They may not have been sophisticated enough to develop an opinion, if that was their task, until they had read at least two documents. Students began by paraphrasing the texts closely and were more likely to reduce information as they read subsequent texts. This tendency to move toward reduction may have been a result of their growing background knowledge.

A second reason that students did not seem to benefit from only reading multiple texts is that they need to be taught what it means to "think like a historian," and that, without this teaching, they will be less able to engage in historical analysis. In other words, students who know more about historical analysis may be more able to engage in it. It is possible, for instance, that the 4 students who exhibited more gisting, evaluating, sourcing, and corroboration may have been more sophisticated readers of historical text, regardless of whether they were familiar with the Tonkin Gulf incident. Textbooks in history are written so that the author's background, stance, and methodology are hidden (Luke, deCastell, & Luke, 1983). Therefore, interpretations of events are presented as fact, not analysis, and two or more interpretations to an event are rarely shared. While original documents and argumentative essays positing different interpretations should help students realize that history is interpretation rather than fact, this idea may be less obvious to students who have relied mainly upon history texts for information and who have been taught to think of history merely as a series of chronicled events.

In addition, students need to be taught to write persuasive essays, with a warrant and supporting evidence. Chambliss' research (1994) shows that there are differences in how students evaluate persuasive essays to formulate their own opinions. Students in the present study made many unsupported statements when asked to form an opinion, even though their notes indicated that they had attended closely to the information in the texts and did have that information at hand. It is possible that these students did not know that they were supposed to provide support for an opinion, even though they clearly learned information that would be appropriate. This is another aspect of the disciplinary knowledge of history, and of other disciplines as well.

A final possible reason for the students' apparent lack of benefit may be their lack of experience in working with multiple texts. As noted earlier, their teacher did not provide such experience, but planned to do so later in the year. Experience (and teacher guidance) may improve students' ability to integrate information from different original source documents.

Author Note. The authors would like to thank Sam Wineburg for his close reading and helpful comments on an earlier version of this text.

References

Alexander, P. A., & Judy, J. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research, 58*, 375–404.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 255–292). White Plains, NY: Longman.

Anderson, T. H., & Armbruster, B. B. (1984). Content area textbooks. In R. C. Anderson, J. Osborn, & R. J. Tierney (Eds.), *Learning to read in American schools* (pp. 193–226). Hillsdale, NJ: Lawrence Erlbaum Associates.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). *Women's ways of knowing*. New York: Basic Books.

Britton, B. K., & Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology, 83*, 329–345.

Chambliss, M. J. (1994). Why do readers fail to change their beliefs after reading persuasive text? In R. G. & P. A. Alexander (Eds.), *Beliefs about text and instruction with text* (pp. 75–89). Hillsdale, NJ: Erlbaum.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Greene, S. (1994). The problems of learning to think like a historian: Writing history in the culture of the classroom. *Educational Psychologist, 29*, 89–96.

Hirsch, E. D. (1987). *Cultural literacy*. Boston: Houghton-Mifflin.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*, 294–303.

Kintsch, W. (1986). Learning from text. *Cognition and Instruction, 3*, 87–108.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363–394.

Leinhardt, G., Beck, I. L., & Stainton, C. (1994). *Teaching and learning in history*. Hillsdale, NJ: Erlbaum.

Luke, C., Decastell, S., & Luke, A. (1983). Beyond criticism: The authority of the school text. *Curriculum Inquiry, 13*, 111–127.

McKeown, M. G., Beck, I. L., & Worthy, M. J. (1993). Grappling with text ideas: Questioning the author. *The Reading Teacher, 46*, 560–566.

Patrick, J. J., & Hawke, S. (1982). Social studies curriculum materials. In *The current state of social studies: A report of Project Span* (pp. 105–185). Boulder, CO: Social Science Education Consortium.

Perfetti, C. A., Britt, M. A., & Georgi, M. C. (1995). *Text-based learning and reasoning: Studies in history*. Hillsdale, NJ: Erlbaum.

Perfetti, C. A., Britt, M. A., Rouet, J.-F., Mason, R. A., & Georgi, M. C. (1993, April). *How students use texts to learn and reason about historical uncertainty*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years*. New York: Holt, Rinehart, & Winston.

Ravitch, D. (1992). *The democracy reader*. New York: Harper-Collins.

Ravitch, D. (Ed.). (1990). *The American reader: Words that moved a nation*. New York: Harper-Collins.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading*

comprehension (pp. 33–58). Hillsdale, NJ: Erlbaum.

Seixes, P. (1993). The community of inquiry as a basis for knowledge and learning: The case of history. *American Educational Research Journal, 30*, 305–324.

Spiro, R. (1980). Constructive processes in prose comprehension and recall. In R. Spiro, B. Bruce, and W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 245–259). Hillsdale, NJ: Erlbaum.

Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1994). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 602–615). Newark, DE: International Reading Association.

Spivey, N. N., & King, J. (1989). Readers as writers composing from sources. *Reading Research Quarterly, 24*, 7–26.

Spoehr, K. T., & Spoehr, L. W. (1994). Learning to think historically. *Educational Psychologist, 29*, 71–77.

Stahl, S. A., & Hynd, C. R. (1994, April). *Selecting history documents: A survey of student reasoning*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Stahl, S. A., Hynd, C. R., Glynn, S., & Carr, M. (1995). Beyond reading to learn: Developing content and disciplinary knowledge through texts. In L. Baker, P. Afflerbach, & D. Reinking (Eds.), *Developing engaged readers in school and home communities* (pp. 139–163). Mahwah, NJ: Erlbaum.

Wineburg, S. S. (1994). Introduction: Out of our past and into our future—the psychological study of learning and teaching history. *Educational Psychologist, 29*, 57–60.

Wineburg, S. S. (1991a). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83*, 73–87.

Wineburg, S. S. (1991b). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal, 28*, 495–519.

Appendix

Analysis Sheet for Notes and Final Product
#25 -- Take side in a debate (Pro-Resolution)

Text	Notes	Final Product
		(1) President Johnson was definitely justified when he asked Congress to pass the Resolution.
1. Gulf of Tonkin Resolution.		
Three days after the second incident in the Tonkin Gulf, the Administration submitted a Joint Resolution to Congress which approved in advance the President's taking "all necessary steps" to assist South Vietnam or any other member or protocol state of the Southeast Asia Treaty Organization.	Administration submitted a Joint Resolution to Congress which approved the President's request to assist the Southeast Asia Treaty Organization	
It was approved unanimously by the House and by a vote of 88 to 2 in the Senate.		(7) The vote passed unanimously in the House of Representives and eighty-two to two in the Senate in approval of military action against the North Vietnamese (Viet Cong).
Whereas naval units of the Communist regime in Vietnam, in violation of the principles of the Charter of the United Nations and of international law,	—Communist Regime violated the Charter	
have deliberately and repeatedly attacked United States naval vessels lawfully present in international waters, and have thereby created a serious threat to international peace;	—attacked UN vessels creating a threat to international peace	(2) He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies (3) and in the process were attacking U.S. ships in the Gulf of Tonkin.
Whereas these attacks are part of a deliberate and systematic campaign of aggression that the Communist regime in North Vietnam has been waging against its neighbors and the nations joined with them in the collective defense of their freedom;	—attacks are part of a deliberate and systematic campaign of aggression that the Communist regime has been waging against its neighbors	(2) He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies

Whereas the United States is assisting the peoples of southeast Asia to protect their freedom and has no territorial, military or political ambitions in that area, but desires only that these peoples should be left in peace to work out their own destinies in their own way:	—US protecting the people's freedom	
Consonant with the Constitution of the United States and the Charter of the United Nations and in accordance with its obligations under the Southeast Asia Collective Defense Treaty, the United States is, therefore, prepared, as the President determines, to take all necessary steps, including the use of armed force, to assist any member or protocol state of the Southeast Asia Collective Defense Treaty requesting assistance in defense of its freedom.	—the US is prepared to take all necessary steps to assist the Southeast Asia Collective Defense Treaty requesting assistance in defense of its freedom	
This resolution shall expire when the President shall determine that the peace and security of the area is reasonably assured by international conditions created by action of the United Nations or otherwise, except that it may be terminated earlier by concurrent resolution of the Congress.	—this power of the President will expire when peace has returned to South Vietnam or as Congress sees fit	
2. The Vote that Congress Can't Forget		
For more than two decades, the Congressional vote that lawmakers most often cite as the one they would like to take back is their 1964 vote for the Gulf of Tonkin Resolution, the resolution that was used as authority for the war in Vietnam. Only two Senators and no Representatives voted no.	The vote that lawmakers would most like to take back is the Gulf of Tonkin Resolution in which two Senators and no Representatives voted no.	(7) The vote passed unanimously in the House of Representives and eighty-two to two in the Senate in approval of military action against the North Vietnamese (Viet Cong).

Twenty-seven of those lawmakers are still in Congress.	—27 of these lawmakers are still in Congress	
And as they prepared for Saturday's vote on what was even more clearly the equivalent of a declaration of war, the resolution to authorize the use of military force in the Persian Gulf, the memories of that earlier vote weighed heavily on the minds of the seven senators and 20 representatives who served then and still serve now.	—They thought about this when making the decision to invade the Persian Gulf.	
He recalled the earlier resolution as one President Johnson had "distorted," and one whose repeal he accomplished as a freshman Senator in 1970.	—the resolution was repealed one year after it was passed	
He added, "The Gulf of Tonkin Resolution was used as a declaration of war and plunged this country in eight or nine years of really disastrous war in Vietnam."	—was used as a declaration of war	
Representative Charles E. Bennett, Democrat of Florida, said, "I am 80 years of age, I have been in this chamber 43 years. Out of the 17,000 votes I have cast, the only one I really regret is the one I cast for the Bay of Tonkin Resolution."	—Charles E. Bennett regrets it.	
Representative Dan Rostenkowski, Democrat of Illinois, did not offer second thoughts about his old vote, but said the Persian Gulf decision was even more difficult for him. "Today's situation is clearer," he said in a statement in the Congressional Record, "The possibility of armed conflict, casualties and even death is much more apparent."	—Persian Gulf situation is clearer	
But few avoided comparisons with Vietnam more generally.	Several lawmakers avoided comparisons of the situations	

When it came time to vote, 12 of the Tonkin veterans voted to authorize force and 14 voted against it.	12 authorized 14 no for the Gulf	(8) Even twenty years after the vote in Congress, twelve Congressmen still believed that the Resolution was justifiable.
The Tonkin Democrats sided against authorizing force.	The Tonkin Democrats sided against authorizing force in the Gulf	
So did Representatives Jack Brooks of Texas, John D. Dingell of Michigan, Dante B. Fascell of Florida, and Jamie L. Whitten of Mississippi, the only current member of Congress who was also on hand in December 1941, to vote the last formal declarations of war, against Japan, Germany and Italy.	James L. Witten of Mississippi, the only current member of Congress who was on hand in December 1941, to vote the last formal declarations of war, against Japan, Germany, & Italy voted Yes.	
3. As I saw it		
He consistently favored strong American involvement, arguing that "aggression" must be stopped. (from "About the Author")	Dean Rusk favored American involvement	
Dean Rusk was Secretary of State under Presidents Kennedy and Johnson, 1961-68. (from "About the Author")	Dean Rusk was Secretary of State under Kennedy and Johnson	
On August 2 and 3, 1964, we received reports that the *USS Maddox* and *USS C. Turner Joy*, American destroyers operating in the Gulf of Tonkin off the coast of North Vietnam, had been attacked by North Vietnamese torpedo boats in two separate incidents.	On August 2 and 3 1964 the *USS Maddox* and *USS C Turner Joy* on the coast of North Vietnam had been attacked	(2) He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies (3) and in the process were attacking U.S. ships in the Gulf of Tonkin.
The Republic of Vietnam today celebrates August 2—the day of the Tonkin Gulf attacks—as part of its national war effort against the Americans, so whatever happened that night in the Tonkin Gulf, evidently it takes credit for it now.	The Republic of Vietnam takes credit for something happening because they celebrate August 2 as part of its efforts against America	

North Vietnam was using coastal waters to infiltrate men and arms into South Vietnam;	North Vietnam was using the coast to infiltrate men and arms into South Vietnam	(2) He saw that North Vietnamese were being hostile toward the South Vietnamese, American allies
South Vietnam under the doctrine of self-defense was trying to block this infiltration and mount retaliatory raids of its own—a secret operation called 34-A, supported by the American Navy. But the destroyers attacked in the Gulf of Tonkin were on intelligence-gathering missions, not participating in South Vietnamese actions along the coast. It is entirely possible that the North Vietnamese thought that our destroyers were involved in these 34-A raids and in blockading operations along North Vietnam's coast to stop their infiltration of the South by sea. But even if Hanoi thought this, it isn't valid to call the exercise of self-defense a provocation.	North Vietnamese could have thought that Americans were part of a South Vietnamese operation called 34-A	(9) The navy needed to protect the South Vietnamese and American intelligence vessels.
		(4) President Johnson could not allow for the continued meaningless destruction of governmental property by North Vietnamese without doing something about it.
Indeed, the Tonkin Gulf Resolution, in which Congress declared its support for the United States' willingness to come to the assistance of those protected by the SEATO Treaty, including the use of armed force "as the President shall determine," was passed rapidly: 88-2 by the Senate and 416-0 by the House.	The resolution was also unanimously voted	(6) Congress also felt he was justified in his actions. (7) The vote passed unanimously in the House of Representives and eighty-two to two in the Senate in approval of military action against the North Vietnamese (Viet Cong).
Some later complained, "We didn't anticipate sending a half million men to South Vietnam," but neither did Lyndon Johnson.		(5) He would not forsee the tragic death and destruction in the future Vietnam War.

	but shortly after people began to change their minds	
I felt the Tonkin Gulf Resolution was not congressional evasion of its war powers responsibility, but an exercise of that responsibility.	Resolution was an exercise of congressional powers not an evasion of them	
		(10) Thus, Lyndon B. Johnson and Congress acted rightly and in good faith regarding the approval of the Gulf of Tonkin Resolution.

NRRC National
Reading Research
Center

318 Aderhold, University of Georgia, Athens, Georgia 30602-7125
3216 J. M. Patterson Building, University of Maryland, College Park, MD 20742